

CLAIMS

We claim:

1. A method of queue congestion control in a multi-stage switch, said multi-stage switch comprising at least one ingress linecard, a plurality of switch elements arranged from a first switch element to a last switch element, and at least one egress linecard, said method comprising:

selectively enqueueing a data packet in an inbound queue in said ingress linecard;

dequeuing said data packet to said first switch element;

enqueueing said data packet in a first switch element queue in said first switch element;

dequeuing said data packet to said last switch element;

enqueueing said data packet in a last switch element queue in said last switch element, wherein said enqueueing comprises determining a last queue status parameter (QSP);

dequeuing said data packet to said egress linecard, said dequeuing further comprising communicating said last QSP to said egress linecard;

enqueueing said data packet in an output queue, said output queue having an output queue status;

wherein said method periodically combines said output queue status with said last QSP, forming a backpressure signal.

2. The method of Claim 1, wherein said selective enqueueing further comprises:

determining an input queue status of said input queue; and

combining said input queue status with said backpressure signal to determine said selectively enqueueing.

3. The method of Claim 1, wherein said last switch element queue comprises a plurality of queue elements and said QSP is based on a number of said queue elements in use in said queue.

4. The method of Claim 1, wherein said last switch element queue
5 comprises a plurality of queue elements and said QSP is based on a fill rate of said plurality of queue elements.

5. The method of Claim 1, wherein said backpressure signal has more than two discrete states.

6. The method of Claim 1, wherein said communicating said
10 backpressure signal uses in-band signaling.

7. The method of Claim 1, wherein said communicating said backpressure signal uses out-of-band signaling.

8. The method of Claim 1, wherein for a plurality of switch elements:
each said enqueueing in a switch element comprises determining a local QSP
15 and combining said local QSP with a prior QSP communicated from a prior switch element to form a new QSP;
each said dequeuing to a next said switch element further comprises communicating said new QSP to said next switch element.

9. The method of Claim 8, wherein said plurality of switch elements
20 comprises all said switch elements in said multi-stage switch.

10. The method of Claim 8, wherein said last switch element queue comprises a plurality of queue elements and said determining a local QSP is based on a number of said queue elements in use in said queue.

5

[illegible]

an ingress linecard comprising circuitry to selectively enqueue a data packet in an inbound queue;

said first switch element, into which said ingress linecard dequeues said data packet, comprising circuitry to enqueue said data packet in a first switch element queue;

said last switch element, into which an upstream switch element dequeues said data packet, comprising circuitry to enqueue said data packet in said last switch element queue and determine a last queue status parameter (QSP); and

an egress linecard, into which said last switch element dequeues said data packet and communicates said last QSP, comprising:

circuitry to enqueue said data packet in an output queue, said output queue having an output queue status;

circuitry to combine said output queue status with said last QSP to periodically form a backpressure signal; and

circuitry to communicate said backpressure signal to said ingress linecard, wherein said circuitry to selectively enqueue said data packet in said ingress linecard is influenced at least in part by said backpressure signal.

13. The apparatus of Claim 12, wherein said circuitry to selectively
30 enqueue said data packet further comprises:
circuitry to determine an input queue status; and

circuitry to combine said input queue status with said backpressure signal to influence said circuitry to selectively enqueue said data packet.

14. The apparatus of Claim 13, wherein said circuitry to combine said input queue status with said backpressure signal further comprises a state machine.

5 15. The apparatus of Claim 12, wherein each said queue comprises a plurality of queue elements and said QSP is based on a number of said queue elements in use in said queue.

10 16. The apparatus of Claim 12, wherein each said queue comprises a plurality of queue elements and said QSP is based on a fill rate of said plurality of queue elements.

17. The apparatus of Claim 12, wherein said backpressure signal has more than two discrete states.

18. The apparatus of Claim 12, wherein said circuitry to communicate said backpressure signal uses in-band signaling.

15 19. The apparatus of Claim 12, wherein said circuitry to communicate said backpressure signal uses out-of-band signaling.

20. The apparatus of Claim 12, having a plurality of switch elements, each of said plurality of switch elements comprising:

20 circuitry to determine a local QSP;
circuitry to combine said local QSP with a prior QSP communicated from a prior switch element to form a new QSP; and
circuitry to communicate said new QSP to said next switch element.

21. The apparatus of Claim 20, wherein said plurality of switch elements comprises all said switch elements in said multi-stage switch.

22. The apparatus of Claim 20, wherein each said queue comprises a plurality of queue elements and said local QSP is determined based on a number of said queue elements in use in said queue.

23. The apparatus of Claim 20, wherein said each said queue comprises a plurality of queue elements and said local QSP is determined based on a fill rate of said plurality of queue elements.

24. The apparatus of Claim 20, wherein said circuitry to combine said local QSP with a prior QSP further comprises a state machine.

25. An apparatus for controlling queue congestion in a multi-stage switch, said multi-stage switch comprising at least one ingress linecard, a plurality of switch elements arranged from a first switch element to a last switch element, and at least one egress linecard, comprising:

an ingress linecard comprising means for selectively enqueueing a data packet in an inbound queue;

said first switch element, into which said ingress linecard dequeues said data packet, comprising means for enqueueing said data packet in a first switch element queue;

said last switch element, into which an upstream switch element dequeues said data packet, comprising means for enqueueing said data packet in said last switch element queue and determine a last queue status parameter (QSP); and

an egress linecard, into which said last switch element dequeues said data packet and communicates said last QSP, comprising:

means for enqueueing said data packet in an output queue, said output queue having an output queue status;

means for combining said output queue status with said last QSP to periodically form a backpressure signal; and

means for communicating said backpressure signal to said ingress linecard, wherein said circuitry to selectively enqueue said data packet in said ingress linecard is influenced at least in part by said backpressure signal.

5

26. A computer system for queue congestion control in a multi-stage switch, said multi-stage switch comprising at least one ingress linecard, a plurality of switch elements arranged from a first switch element to a last switch element, and at least one egress linecard; said computer system comprising computer instructions for:

10 selectively enqueueing a data packet in an inbound queue in said ingress linecard;

dequeuing said data packet to said first switch element;

enqueueing said data packet in a first switch element queue in said first switch

15 element;

dequeuing said data packet to said last switch element;

enqueueing said data packet in a last switch element queue in said last switch element, wherein said enqueueing comprises determining a last queue status parameter (QSP);

20 dequeuing said data packet to said egress linecard, said dequeuing further comprising communicating said last QSP to said egress linecard;

enqueueing said data packet in an output queue, said output queue having an output queue status;

wherein said method periodically combines said output queue status with said last

25 QSP, forming a backpressure signal.

27. A computer-readable storage medium, comprising computer instructions for queue congestion control in a multi-stage switch, said multi-stage switch comprising at least one ingress linecard, a plurality of switch elements arranged from a first switch element to a last switch element, and at least one egress

5 linecard; said computer instructions comprising:

selectively enqueueing a data packet in an inbound queue in said ingress linecard;

dequeuing said data packet to said first switch element;

10 enqueueing said data packet in a first switch element queue in said first switch element;

dequeuing said data packet to said last switch element;

enqueueing said data packet in a last switch element queue in said last switch element, wherein said enqueueing comprises determining a last queue status parameter (QSP);

15 dequeueing said data packet to said egress linecard, said dequeuing further comprising communicating said last QSP to said egress linecard;

enqueueing said data packet in an output queue, said output queue having an output queue status;

20 wherein said method periodically combines said output queue status with said last QSP, forming a backpressure signal.

28. A computer data signal embodied in a carrier wave, comprising computer instructions for queue congestion control in a multi-stage switch, said multi-stage switch comprising at least one ingress linecard, a plurality of switch elements arranged from a first switch element to a last switch element, and at least one egress linecard; said computer instructions comprising:

25 selectively enqueueing a data packet in an inbound queue in said ingress linecard;

30 dequeueing said data packet to said first switch element;

enqueueing said data packet in a first switch element queue in said first switch element;

dequeueing said data packet to said last switch element;

enqueueing said data packet in a last switch element queue in said last switch element, wherein said enqueueing comprises determining a last queue status parameter (QSP);

dequeueing said data packet to said egress linecard, said dequeueing further comprising communicating said last QSP to said egress linecard;

enqueueing said data packet in an output queue, said output queue having an output queue status;

wherein said method periodically combines said output queue status with said last QSP, forming a backpressure signal.